

WHAT IS CLAIMED IS:

1. In an information processing system, a method for recognizing speech to be recognized, the method comprising the steps of:

maintaining a model of speech accent that is established based on training speech data, wherein the training speech data includes at least a first set of training speech data, and wherein establishing the model of speech accent includes not using any phone or phone-class transcription of the first set of training speech data;

deriving features from the speech to be recognized, the features hereinafter referred to as features for identifying accent;

identifying accent of the speech to be recognized based on the features for identifying accent and on the model of speech accent; and

recognizing the speech to be recognized based at least in part on the identified accent of the speech.

2. The method of claim 1, wherein the establishing the model of speech accent includes estimating model parameters using known accent of the first set of training speech data.

3. The method of claim 2, wherein the known accent of the first set of speech training data includes mandarin Chinese.

4. The method of claim 2, wherein the known accent of the first set of speech training data includes Cantonese Chinese.

5. The method of claim 1, wherein the model of speech accent includes a hidden Markov model trained to model an accent without states that specifically model predetermined phones or classes of phones.

6. The method of claim 1, wherein the step of recognizing the speech to be recognized based at least in part on the identified accent of the speech comprises:

deriving features, hereinafter referred to as features for recognizing speech, from the speech to be recognized; and

evaluating the features for recognizing speech using at least a speech recognition model that is deemed appropriate for the identified accent.

7. The method of claim 6, wherein the features for recognizing speech are not identical with the features for identifying accent.

8. The method of claim 7, wherein the features for identifying accent are reduced from a larger dimension of possible features.

9. The method of claim 8, wherein the features for identifying accent are reduced from a larger dimension of possible features using eigenvalue decomposition.

10. The method of claim 8, wherein the features for identifying accent are reduced from a larger dimension of possible features by determining and dropping less-useful possible features during training.

11. The method of claim 6, wherein the speech recognition model that is deemed appropriate for the identified accent includes an acoustic model that has been adapted for the identified accent.

12. The method of claim 11, wherein the acoustic model that has been adapted for the identified accent was adapted without using accented training speech data.

13. The method of claim 11, wherein the acoustic model that has been adapted for the identified accent was adapted using training speech data of a language, other than language of the speech to be recognized, that is associated with the identified accent.

14. The method of claim 13, wherein the language of the speech to be recognized is English, and the language that is associated with the identified accent is

mandarin Chinese if the identified accent is a mandarin Chinese accent.

15. In an information processing system, a method for recognizing speech to be recognized, the method comprising the steps of:

identifying accent of the speech to be recognized based on information derived from the speech to be recognized; and

evaluating features derived from the speech to be recognized using at least an acoustic model that has been adapted for the identified accent using training speech data from a language, other than primary language of the speech to be recognized, that is associated with the identified accent.

16. The method of claim 15, wherein the language of the speech to be recognized is English, and the language that is associated with the identified accent is mandarin Chinese if the identified accent is a mandarin Chinese accent.

17. The method of claim 16, wherein the language that is associated with the identified accent is Cantonese Chinese if the identified accent is a Cantonese Chinese accent.

18. The method of claim 15, wherein adapting the acoustic model that has been adapted included transforming phonetic transcriptions of the training speech data, from the language that is associated with the identified accent, into phonetic transcriptions according to the language of the speech to be recognized, and then using the result as if it were training speech data of accented speech for model adaptation.

19. A system for recognizing speech to be recognized, the system comprising:

an accent identifier that is configured to identify accent of the speech to be recognized, wherein the accent identifier comprises a model of speech accent that is established based at least in part on using certain training speech data without using any phone or phone-class transcription of the certain training speech data; and

a recognizer that is configured to use models, including a model deemed appropriate for the accent identified by the accent identifier, to recognize the speech to be recognized.

20. The system of claim 19, wherein the model of speech accent is established based at least in part on using the certain training speech data and using known accent of the certain training speech data.

21. The system of claim 20, wherein the certain speech training data includes mandarin Chinese-accented training data.

22. The system of claim 21, wherein the certain speech training data further includes Cantonese Chinese-accented training data.

23. The system of claim 19, wherein the model of speech accent includes a hidden Markov model trained to model an accent and not predetermined individual phones or classes of phones.

24. The system of claim 19, wherein the accent identifier comprises an analyzer that derives features from the speech to be recognized, and the features are features that have been reduced from a larger dimension of possible features.

25. The system of claim 24, wherein the features have been reduced from a larger dimension of possible features using eigenvalue decomposition.

26. The system of claim 25, wherein the features have been reduced from a larger dimension of possible features by determining and dropping less-useful possible features during training.

27. The system of claim 19, wherein the model that is deemed appropriate for the identified accent includes an acoustic model that has been adapted for the identified

accent.

28. The system of claim 27, wherein the acoustic model that has been adapted for the identified accent was adapted without using accented training data.

29. The system of claim 27, wherein the acoustic model that has been adapted for the identified accent was adapted using training data from a language, other than primary language of the speech to be recognized, that is associated with the identified accent.

30. The system of claim 29, wherein the language of the speech to be recognized is English, and the language that is associated with the identified accent is mandarin Chinese if the identified accent is a mandarin Chinese accent, and the language that is associated with the identified accent is Cantonese Chinese if the identified accent is a Cantonese Chinese accent.

31. A system for recognizing speech to be recognized, the system comprising:
an accent identification module that is configured to identify accent of the speech to be recognized; and
a recognizer that is configured to use models to recognize the speech to be recognized, wherein the models include at least an acoustic model that has been adapted for the identified accent using training speech data of a language, other than primary language of the speech to be recognized, that is associated with the identified accent.

32. The system of claim 31, wherein the language of the speech to be recognized is English, and the language that is associated with the identified accent is mandarin Chinese if the identified accent is a mandarin Chinese accent.

33. The system of claim 32, wherein the language that is associated with the identified accent is Cantonese Chinese if the identified accent is a Cantonese Chinese accent.

